

Figure 1 - Computer System

(With Connections to Local Area Network, Local Storage, and Internet)

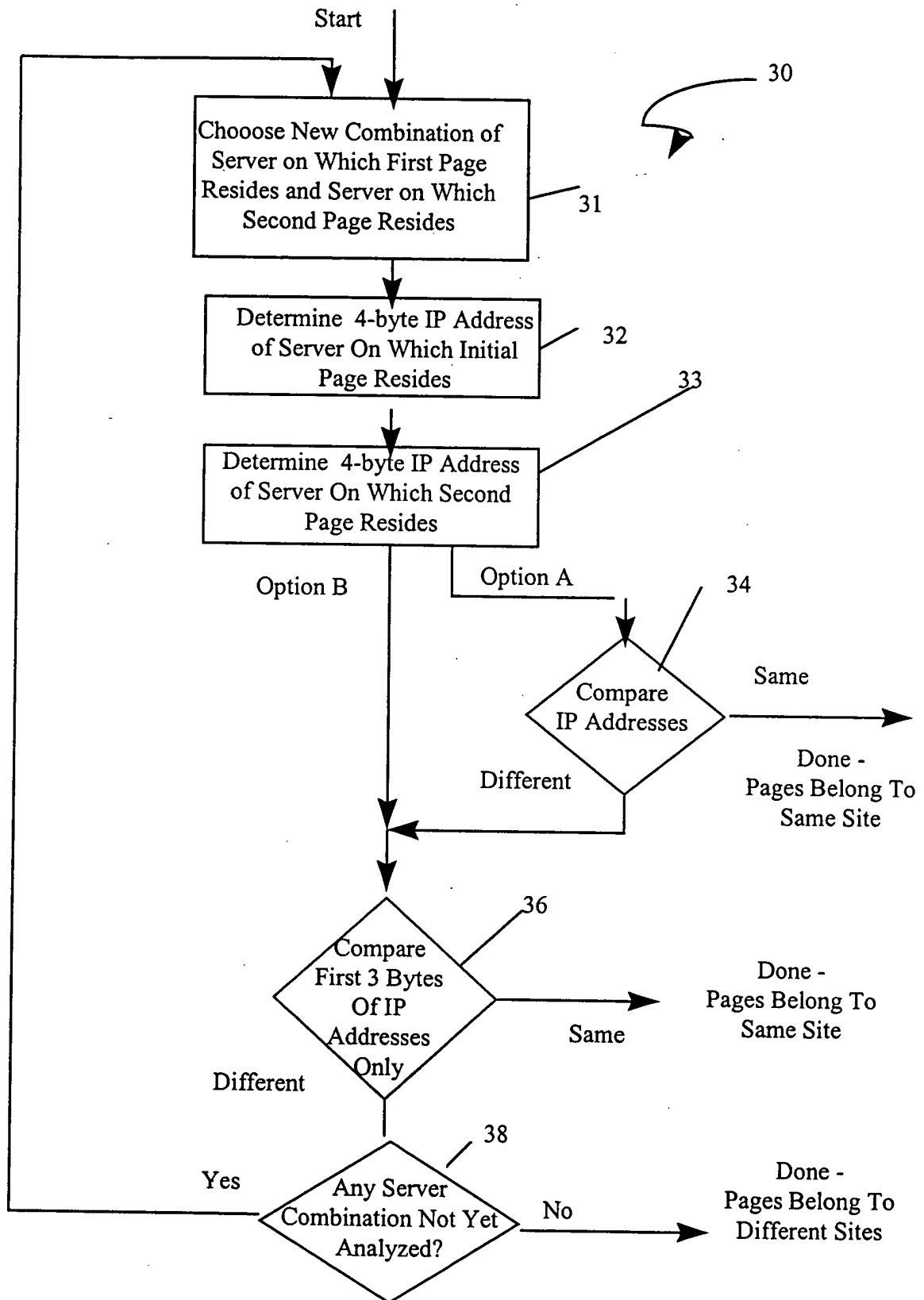


Figure 2 - Process to Determine If Two Linked Web Pages Belong To Same Web Site

Boundary of
Site After
Process
Completed

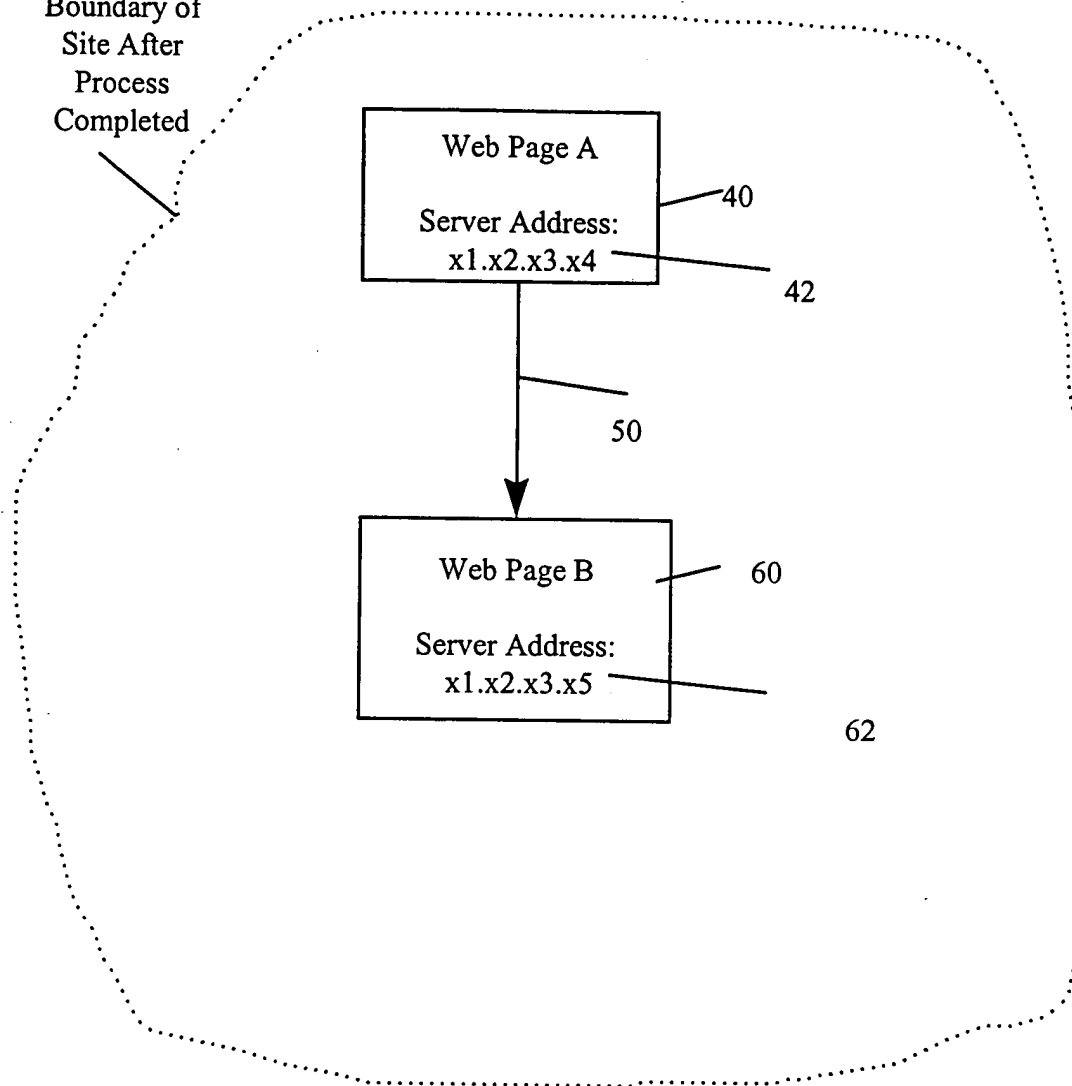


Figure 3 - Two Linked Web Pages

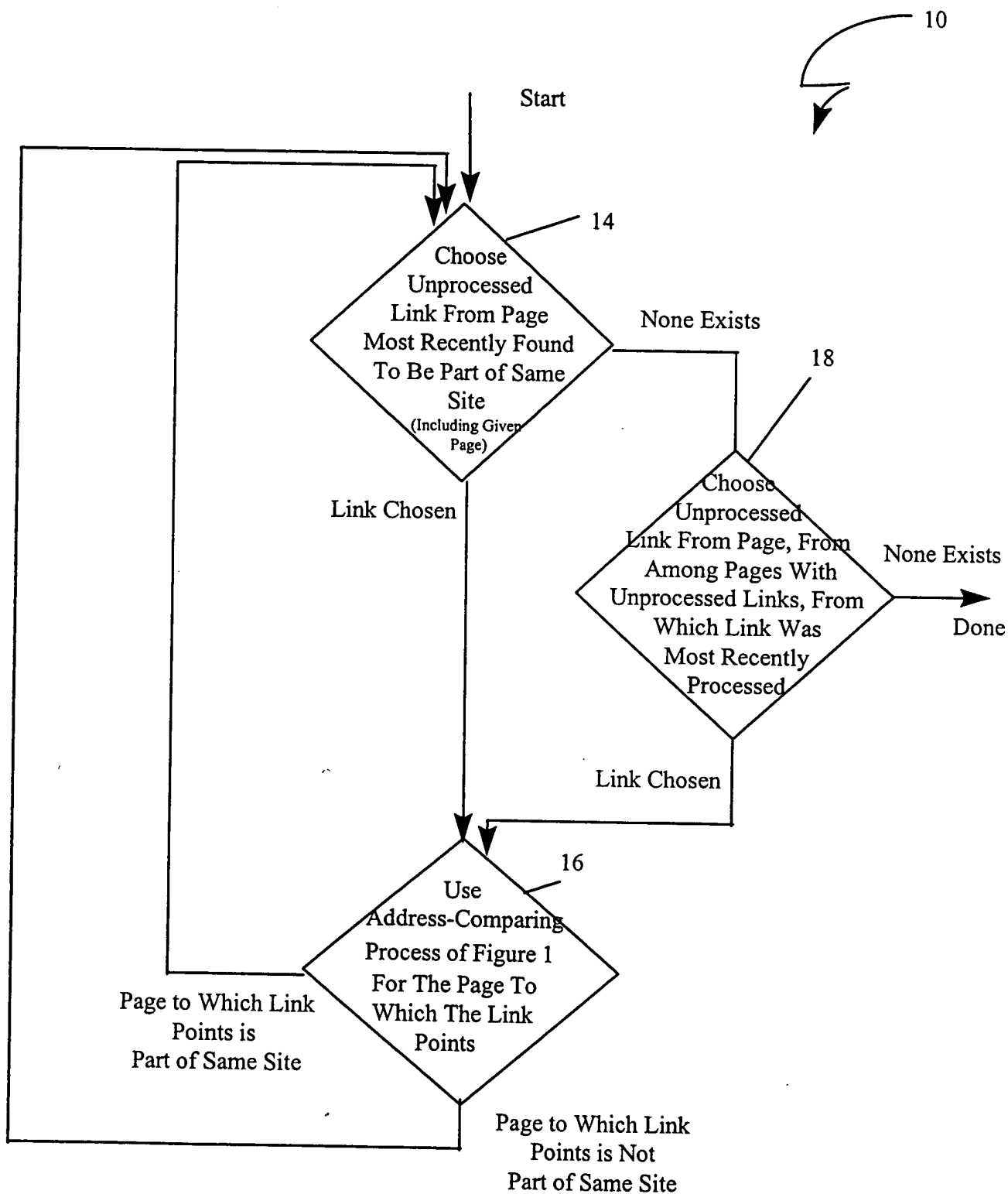


Figure 4 - Site Mapping Process

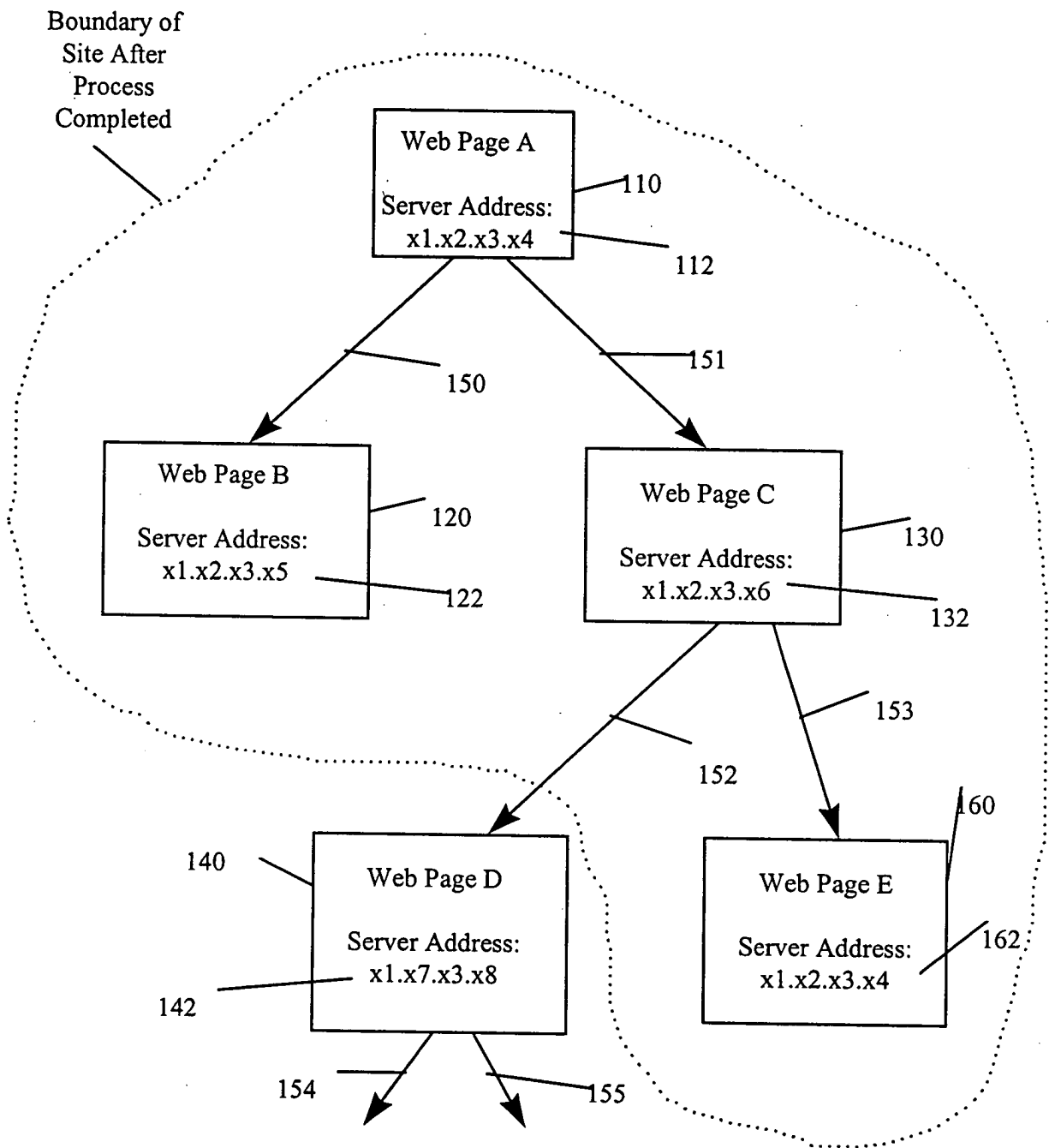


Figure 5 - Multiple Linked Web Pages

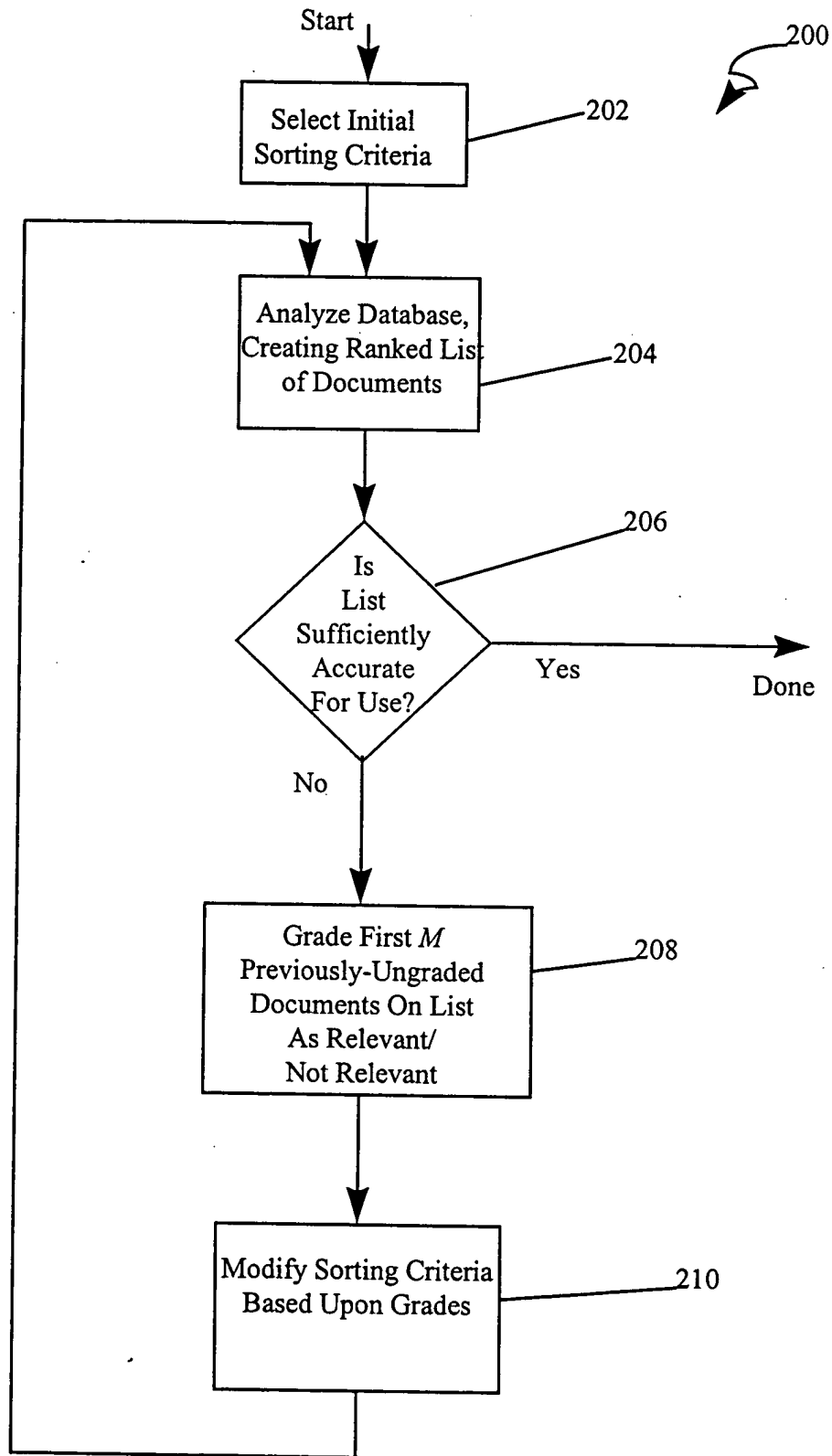


Figure 6 - Process to Classify By IR Techniques

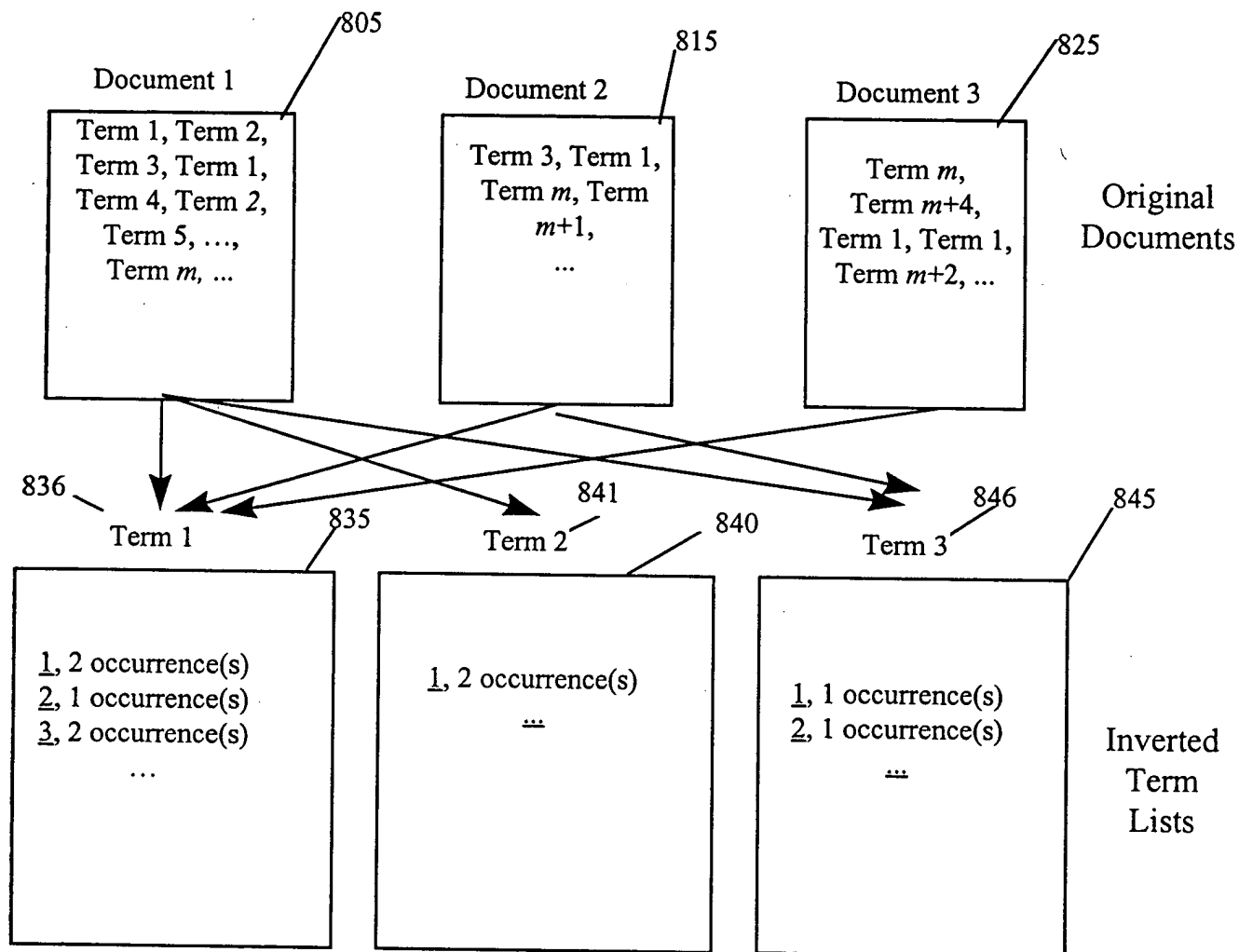


Figure 7 - Relationship Between Documents and Inverted Term Lists (Prior Art)

Document #	Document URL
1	URL 1
2	URL 2
3	URL 3
...	

Lookup
Table

Term #	Term	ITL Address	Number of Documents Containing Term	TF _{MAX}
1	"Term 1"	Loc. 1	N ₁	TF _{M1}
2	"Term 2"	Loc. 2	N ₂	TF _{M2}
3	"Term 3"	Loc. 3	N ₃	TF _{M3}

Lookup
Table

Figure 8 - Lookup Tables Which May Be Used With Inverted Term Lists (Prior Art)

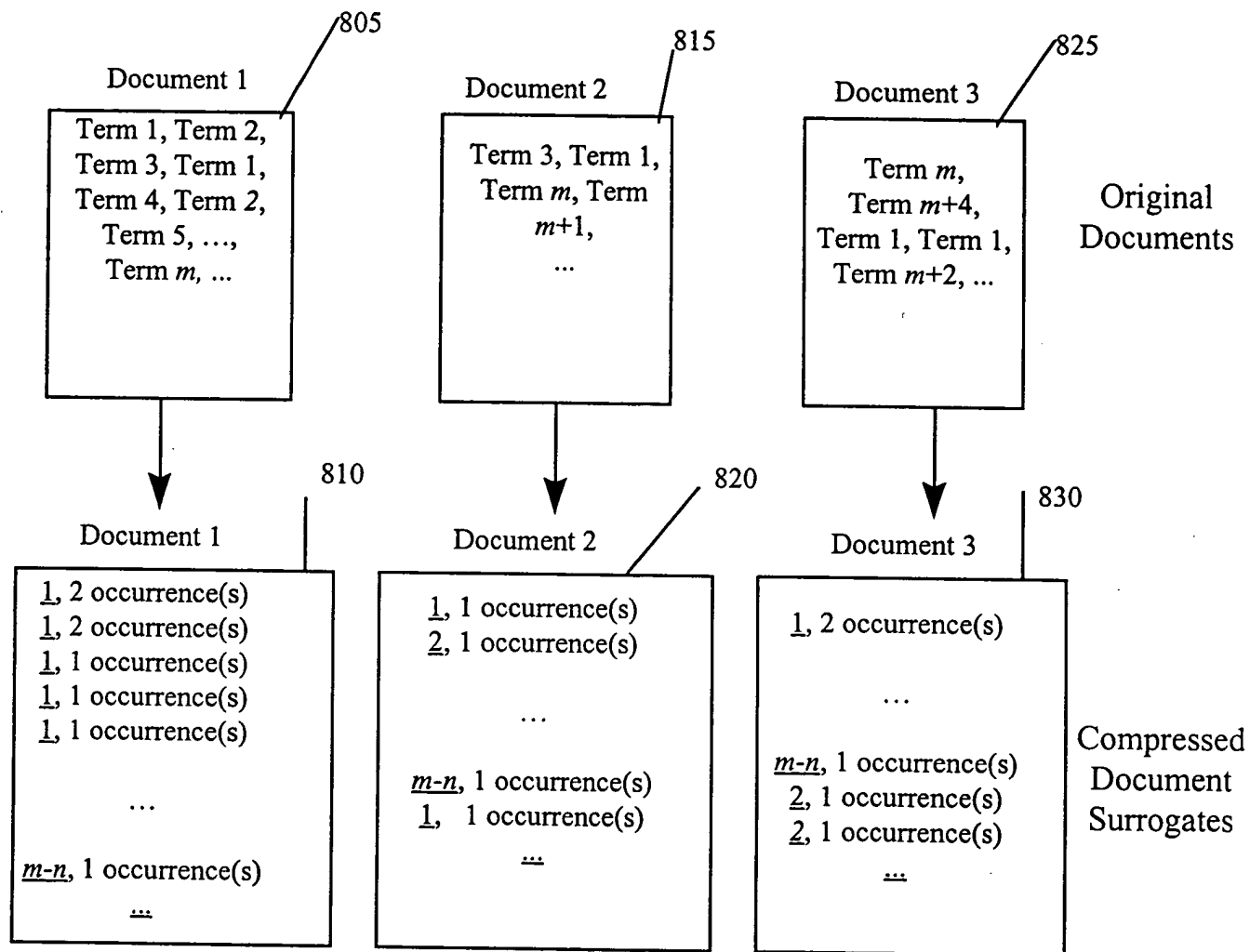


Figure 9 - Relationship Between Documents and Compressed Document Surrogates

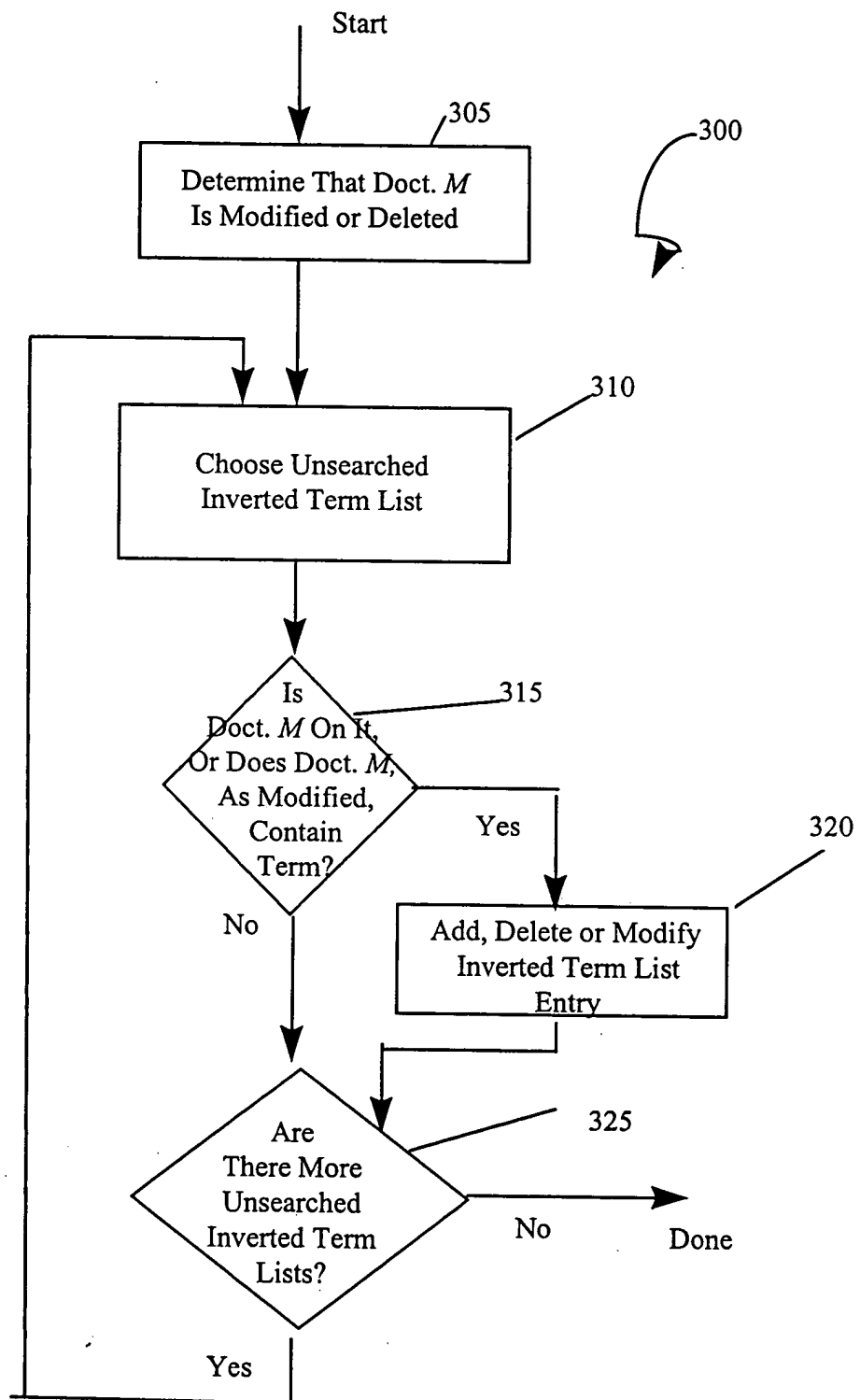


Figure 10 - Database Modification Process - Conventional Technique (Prior Art)

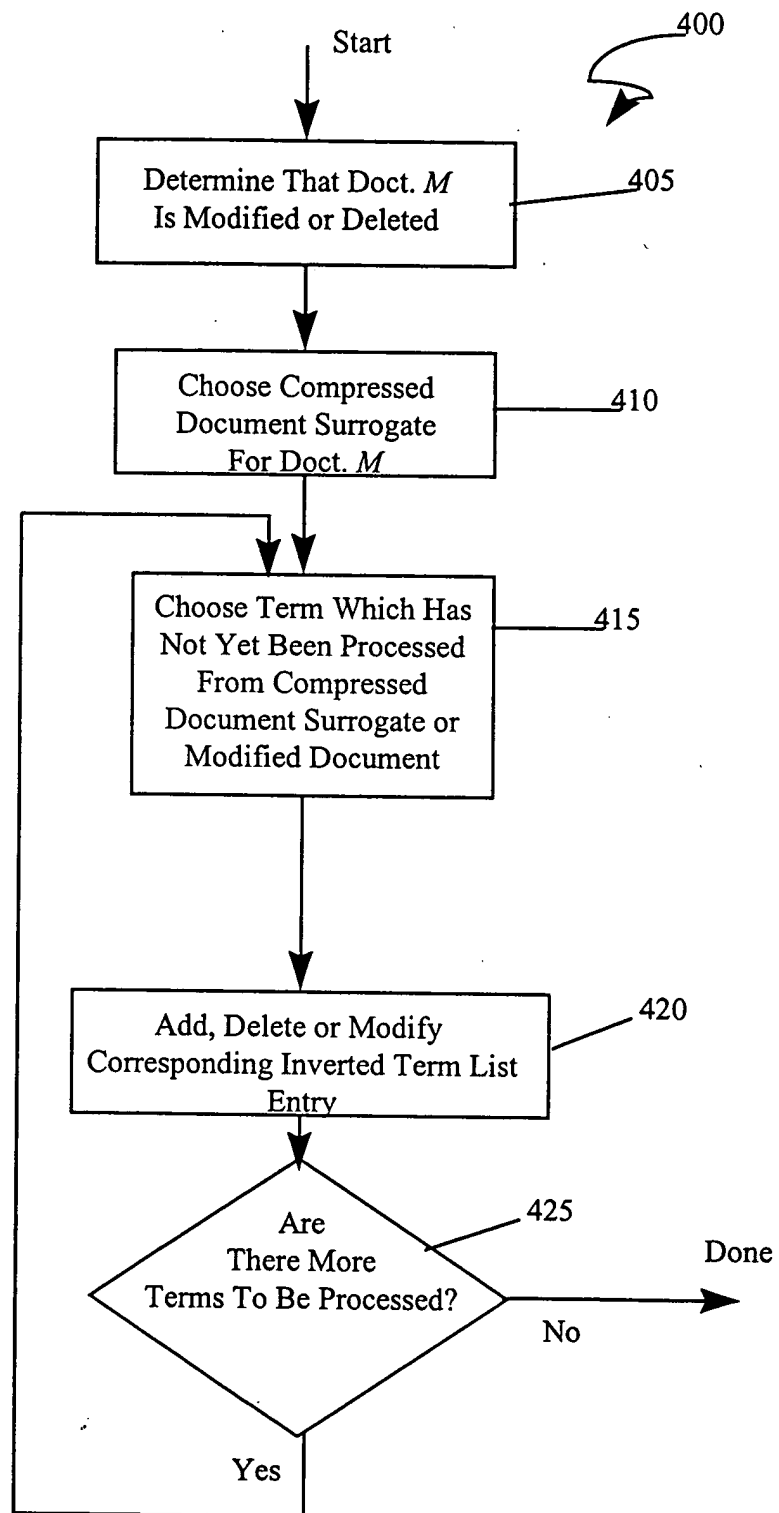


Figure 11 - Database Modification Process - Compressed Document Surrogate Technique

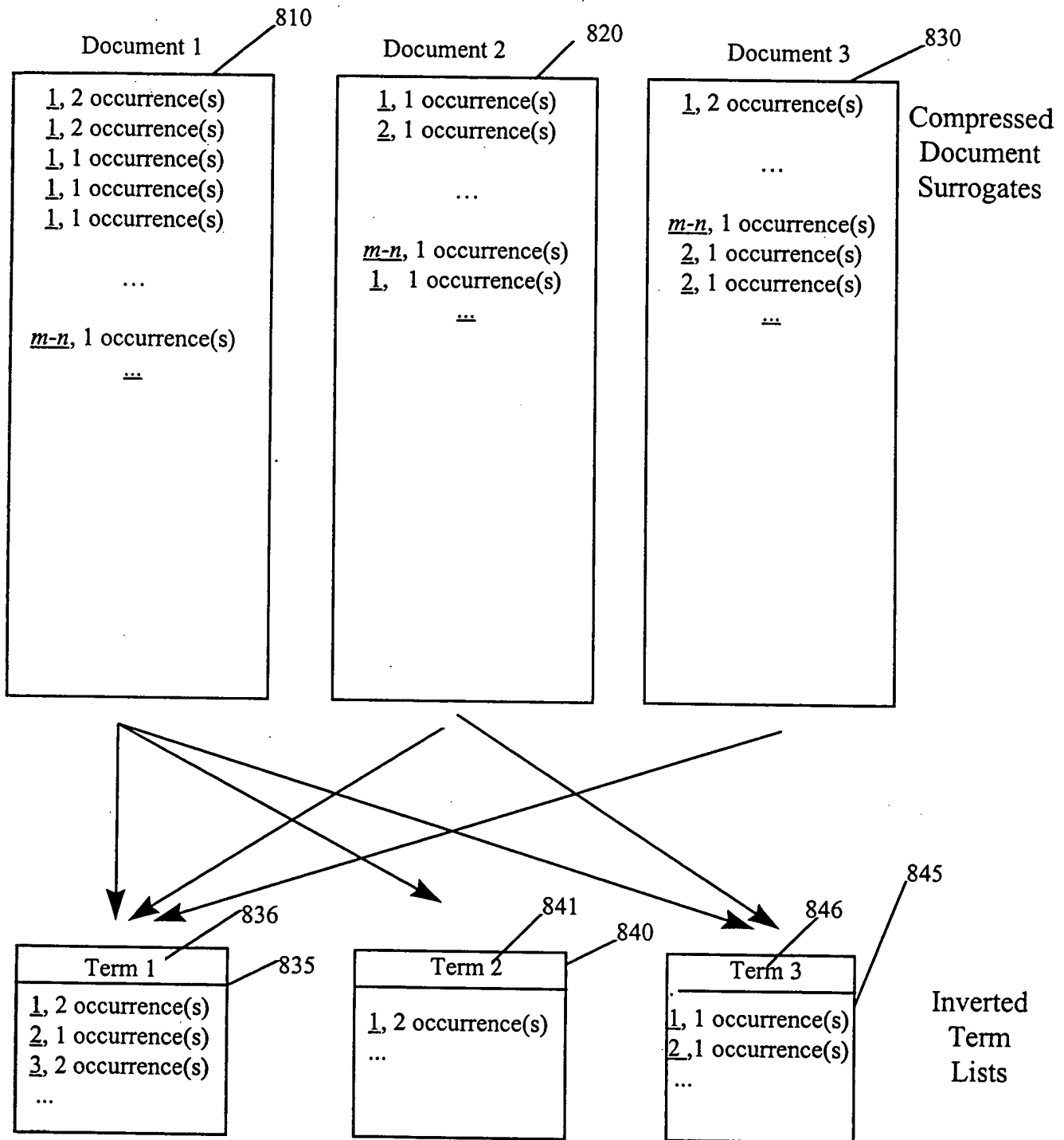


Figure 12 - Relationship Between Compressed Document Surrogates and Inverted Term Lists

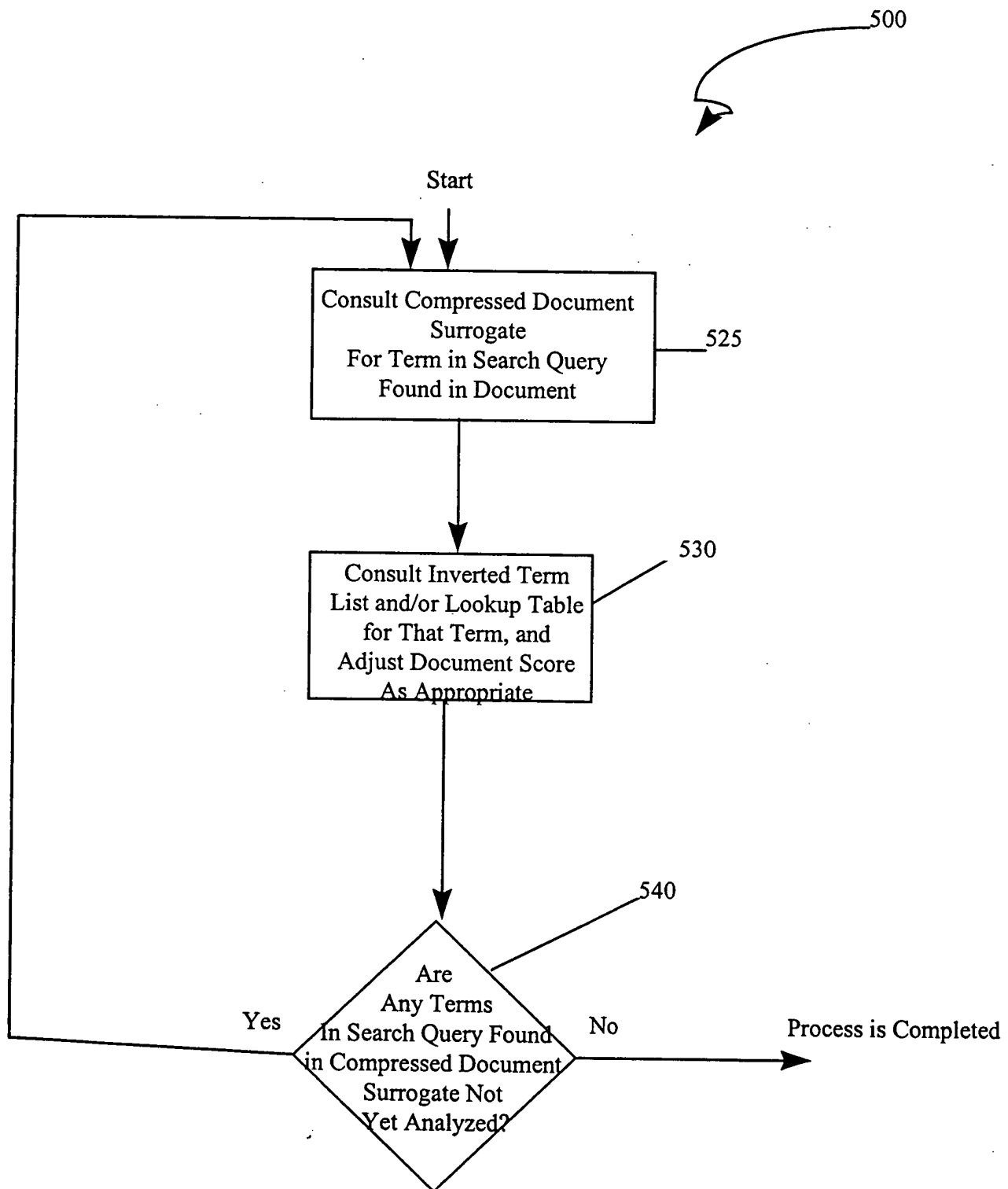


Figure 13 - Process of Calculating Document Score Using Compressed Document Surrogate

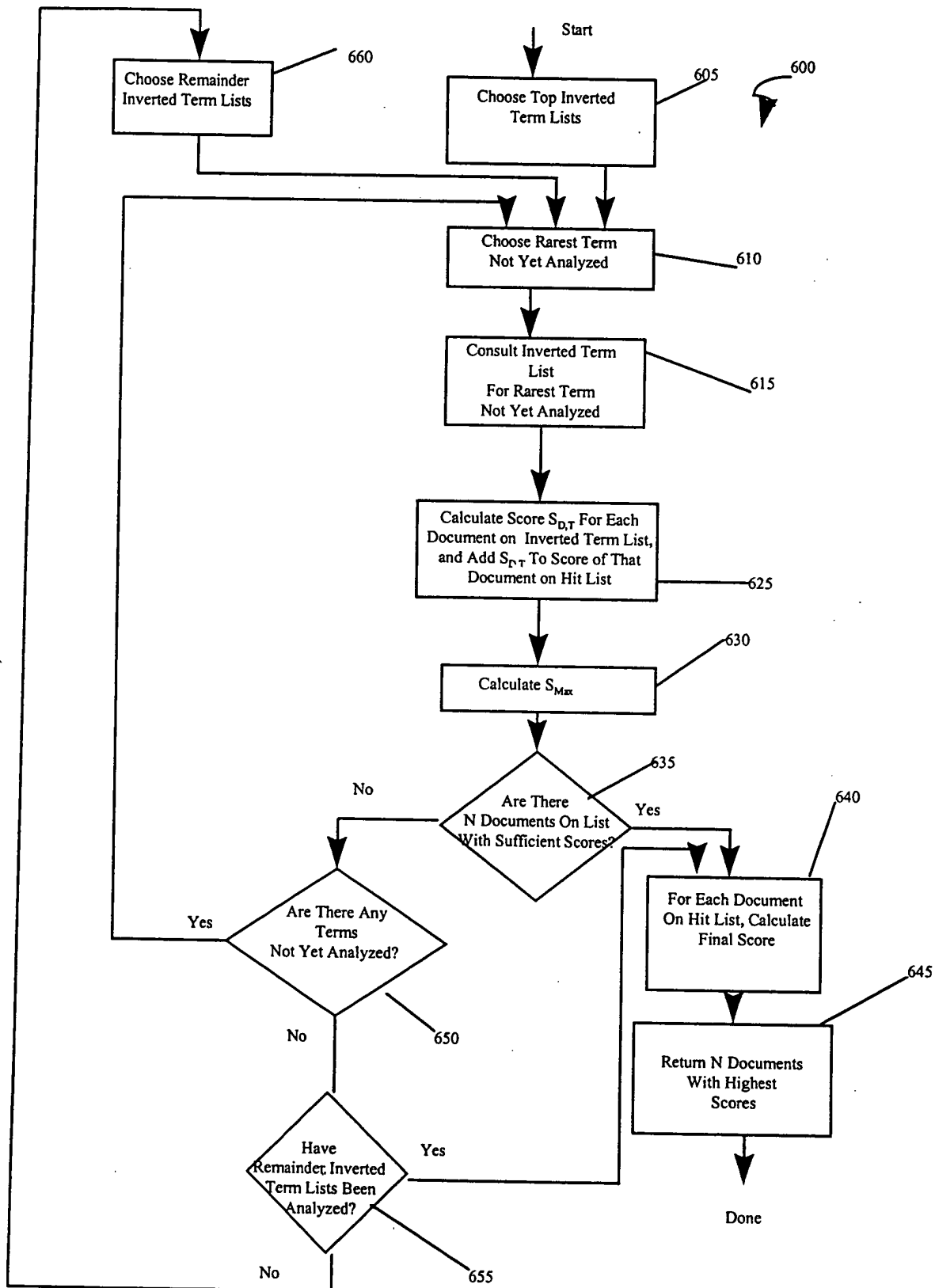


Figure 14 - Process of Carrying Out Search Query Using Compressed Document Surrogates